Error Analysis and Adaptive Localization for Ensemble Methods in Data Assimilation

África Periáñez, Hendrik Reich, Roland Potthast

DWD, Germany

Offenbach, March 6, 2013

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Table of contents

Introduction
- Challenge
- Introduction
- Ensemble Kalman Filter

Error Analysis on Ensemble Methods
- Error analysis with and without background error
- Localization

Numerical results
- Introduction
- Results
- Localization

Challenge

- Understand the basic properties of localization in the ensemble Kalman filter scheme.
- Define an adaptive localization depending on the density of data, observation and background error.
- Decomposition of the error sources to determine its effect on the optimal localization length scale.
- Perspective of approximation theory and functional analysis.
- Addressed with numerical experimental results.

Introduction

In order to find out φ we should minimize the functional

$$J(\varphi) := \left\| \varphi - \varphi^{(b)} \right\|^2 + \left\| f - H\varphi^{(b)} \right\|^2.$$

The normal equations are obtained from first order optimality conditions

$$\nabla_{\varphi} J = 0.$$

Usually, the relation between variables at different points is incorporated by using covariances/weighted norms:

$$J(\varphi) := \| \varphi - \varphi^{(b)} \|_{B^{-1}}^2 + \| f - H\varphi \|_{R^{-1}}^2,$$

The update formula is now

$$\varphi^{(a)} = \varphi^{(b)} + BH^*(R + HBH^*)^{-1}(f - H\varphi^{(b)})$$

Ensemble Kalman Filter

- In the KF method B evolves with the model dynamics: $B_{k+1} = MB_kM^*$.
- EnKF[1] is a Monte Carlo approximation to the KF.
- EnKF methods use reduced rank estimation techniques to aproximate the classical filters.
- The ensemble matrix $Q_k := \left(\varphi_k^{(1)} - \overline{\varphi}_k^{(b)}, ..., \varphi_k^{(L)} - \overline{\varphi}_k^{(b)}\right)$.
- In the EnKF methods the background convariance matrix is represented by $B := \frac{1}{L-1} Q_k Q_k^*$.
- Update solved in a low-dimensional subspace

$$U^{(L)} := \text{span}\{\varphi_k^{(1)} - \overline{\varphi}_k^{(b)}, ..., \varphi_k^{(L)} - \overline{\varphi}_k^{(b)}\}.$$

[1] Evensen 1994

- The updates of the EnKF are a linear combination of the columns of Q_k.

$$\varphi_k - \varphi_k^{(b)} = \sum_{l=1}^{L} \gamma_l \left(\varphi_k^{(l)} - \overline{\varphi}_k^{(b)} \right) = Q_k \gamma$$

$$\varphi_k^{(a)} = \varphi_k^{(b)} + Q_k Q_k^* H^* (R + H Q_k Q_k^* H^*)^{-1} (f_k - H\varphi_k^{(b)})$$

- The previous cost function

$$J(\varphi) := \|\varphi - \varphi^{(b)}\|_{B^{-1}}^2 + \|f - H\varphi\|_{R^{-1}}^2$$

results now in this expresion to minimize:

$$J(\gamma) := \|Q_k \gamma\|_{B_k^{-1}}^2 + \|f_k - H\varphi_k^{(b)} - HQ_k\gamma\|_{R^{-1}}^2$$

- We denote the analysis error $E_k := \|\varphi^{(a)} - \varphi^{(true)}\|$

Error analysis without background contribution

Lemma

Assume that H is injective, that we study true measurement data $f = H\varphi^{(true)}$ *and consider the EnKF with data term only*

$$J^{(data)}(\gamma) = \|(f - H\varphi^{(b)}) - HQ_k\gamma\|^2_{R^{-1}}$$

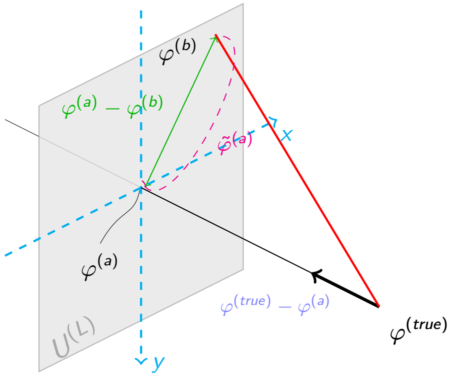
Then, for the analysis $\varphi^{(a)}$ *calculated by the EnKF the difference* $\varphi^{(a)} - \varphi^{(b)}$ *is the orthogonal projection of* $\varphi^{(true)} - \varphi^{(b)}$ *onto the ensemble space* $U_k^{(L)}$ *and the analysis error is given by* [2]

$$E_k = d_{H^* R^{-1} H}\Big(U_k^{(L)}, \varphi_k^{(true)} - \varphi^{(b)}\Big),$$

where the right-hand side denotes the distance between a point $\psi = \varphi_k^{(true)} - \varphi^{(b)}$ *and the subspace* $U^{(L)}$ *with respect to the norm induced by the scalar product* $< ., . >_{H^* R^{-1} H}$.

[2] Proof in Perianez A., Reich H. and Potthast R. *In preparation.*

Illustration of Lemma



Error analysis with background term

Theorem

For H injective, the analysis $\tilde{\varphi}^{(a)}$ generated by the minimization of the whole cost function within the EnKF for perfect data $f^{(true)}$ satisfies the estimate

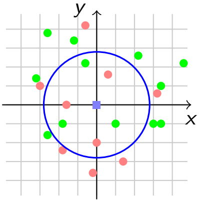
$$\begin{aligned} \left\| \tilde{\varphi}^{(a)} - \varphi^{(true)} \right\|_{HR^{-1}H} &\leq \sqrt{q^2 (E^{(b)})^2 + (1-q^2) E_{min}^2} \\ &= E^{(b)} \sqrt{q^2 + (1-q^2) \frac{E_{min}^2}{(E^{(b)})^2}} \end{aligned}$$

with some constant $q < 1$ depending on B, R and H, where

$$\begin{aligned} E_{min} &:= \min_{\varphi \in U^{(L)}} \left\| \varphi - \varphi^{(true)} \right\|_{H^* R^{-1} H} = \left\| \check{\varphi}^{(a)} - \varphi^{(true)} \right\|_{H^* R^{-1} H}, \\ E^{(b)} &:= \left\| \varphi^{(b)} - \varphi^{(true)} \right\|_{H^* R^{-1} H}. \end{aligned}$$

Localization

- Localization denotes the restriction to a subset of physical space.
- Study the analysis in dependence of the localization radius ρ when the domain D is given by a ball $D = B_\rho(x_0)$.



- Localization function χ_ρ depending on ρ such that[3]

$$\chi_\rho(x) := \begin{cases} \chi_\rho(x) & x \in D \\ 0 & \text{otherwise.} \end{cases}$$

- R localization[4] modifies the observation error covariance matrix to suppress the influences of distant observations $\rightarrow R_{loc} := \chi \cdot R$

$$E^{(\rho)}(x_0) := \left\| \breve{\varphi}^{(a,\rho)} - \varphi^{(true,\rho)} \right\|_{H^* R_{loc}^{-1} H}$$

[3] Houtekamer et al. 1998

[4] Hunt et al. 2007

Localization. Convergence results.

Theorem

We study assimilation in the case where true data $\varphi^{(true)}$ *are used and* $\hat{\varphi}^{(a,\rho)}$ *is chosen such that* $\hat{\varphi}^{(a,\rho)} - \varphi^{(b)}$ *is the orthogonal projection of* $\varphi^{(true)} - \varphi^{(b)}$ *onto the ensemble space* $U^{(L)}$*. Assume that there is* c*,* $C > 0$ *such that for all* $x \in D$ *there is* $l \in \{1, ..., L\}$ *such that*

$$|\varphi^{(l)}(x) - \varphi^{(b)}(x)| \geq c,$$

and that the ensemble members are continuously differentiable on D *with*

$$|\nabla(\varphi^{(j)}(x) - \varphi^{(b)}(x))| \leq C, \quad x \in D, \quad j \in \{1, ..., L\}.$$

Further assume that $\varphi^{(true)} - \varphi^{(b)}$ *is continuously differentiable on* D*. Then, we have*

$$\sup_{x_0 \in D} E^{\rho}(x_0) \leq \tilde{C}\rho \quad \to 0, \quad \rho \to 0$$

with some constant $\tilde{C}$ *depending on* C*,* H *and* R*.*

Remarks

- Using last two theorems we can derive

$$\left\|\tilde{\varphi}^{(a,\rho)} - \varphi^{(true,\rho)}\right\|_{H^* R_{loc}^{-1} H} \leq \sqrt{E^{(b,\rho)} q^2 + (1-q^2) C \rho^2}.$$

- The first term q^2 in the square root reflects the influence of the *background error*.
- The second *approximation error* term can be made small by reducing the localization radius ρ.
- In a balanced relationship between background and data, q is between 0 and 1.
- *Effective observation error*: With data error, ρ needs to be kept sufficiently large since it also controls the number of observations used for the assimilation.

A one-dimensional example

- Ensemble space given by linear functions
  $U^{(L)} := \{a + bx : \ \ a, b \in \mathbb{R}\} \subset L^2([0, A])$.
- The truth $\varphi^{(true)}$ given by a quadratic function
  $\varphi^{(true)}(x) := B \cdot (x - C)^2, \ \ x \in [0, A]$.
- Observations from a Gaussian distribution with variance σ_{obs}.
- Localization by a decomposition of $[0, A]$ into $q \in \mathbb{N}$ subsets
  $[A_j, A_{j+1}]$ where $A_j := \frac{j \cdot A}{q}, \ \ j = 0, ..., q$.
- Localization radius here $\rho = A/2q$.
- On each subset the analysis is carried out by solving the least squares problem in $U^{(L)}|_{[A_j, A_{j+1}]}$.

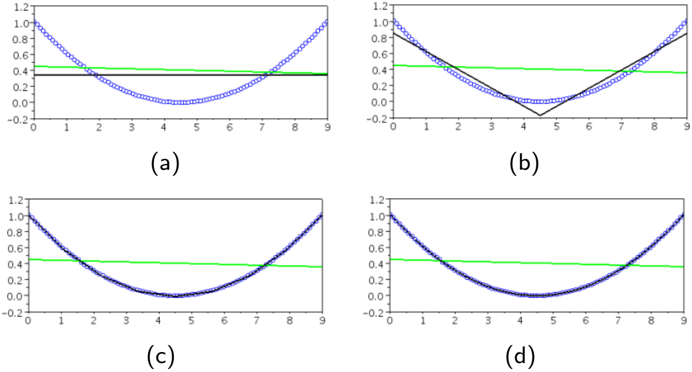


Figure: The truth (blue line) overlaps with the observations (blue circles) due to σ_{obs} takes very small values ($\sigma_{obs} = 0.0005$). The green line shows background information. In (a) we show the analysis without any localization. Localization radii gradually decreases in (b), (c), and (d).

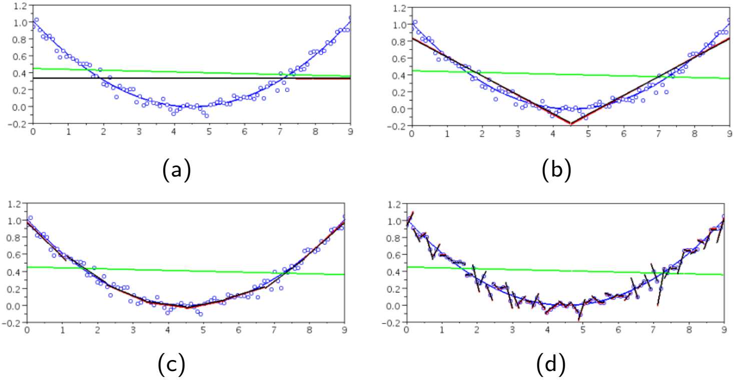


Figure: Observation error $\sigma_{obs} = 0.05$. No localization is applied for (a). In (b), (c) and (d) the localization radii is progressively reduced.

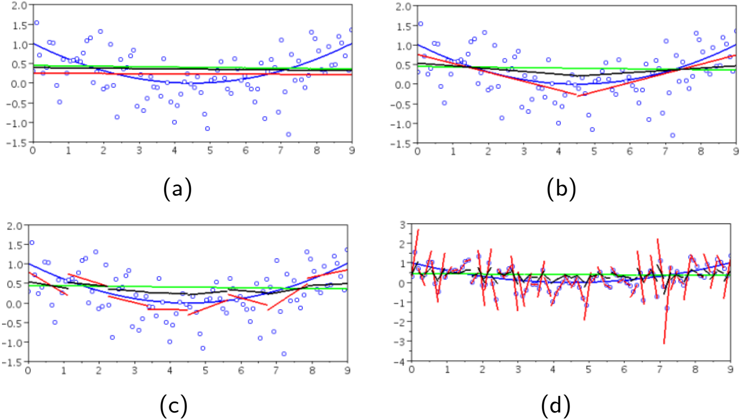


Figure: Higher observation error, $\sigma_{obs} = 0.5$, is provided. The analsyis in (a) is computed with no localization, being progressively smaller in (b), (c) and (d) cases.

Estimating fronts in a 2d example with LETKF [5]

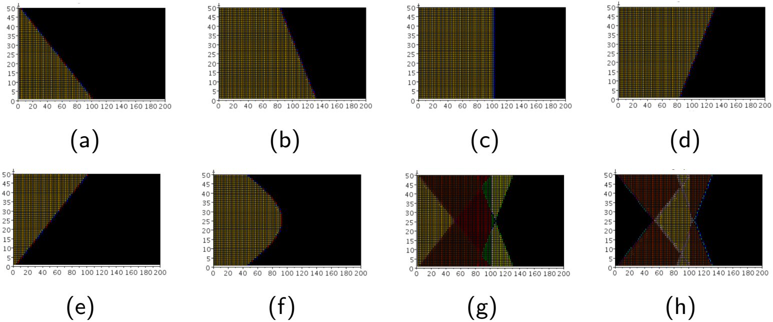


Figure: The different ensemble members are shown in (a)-(e). The truth is displayed in (f). The first guess mean and spread are plotted in (g) and (h), respectively.

[5] Hunt et al. 2007

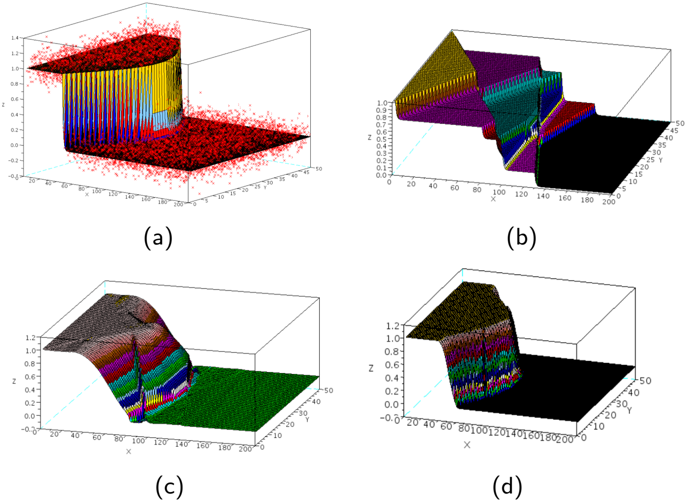
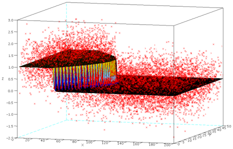
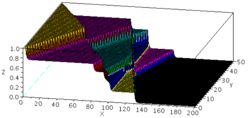


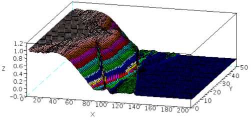
Figure: Truth (front) and observations (red crosses) for $\sigma_{obs} = 0.1$ in (a). Truth (a) is approximated without any localization procedure in (b). In (c) $\rho = 15$, and in (d) $\rho = 5$.



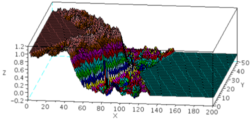
(a)



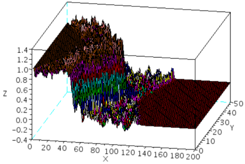
(b)



(c)



(d)



(e)

Figure: Truth and observations for $\sigma_{obs} = 0.5$. (a) is approximated without any localization in (b), with $\rho = 15$ in (c), $\rho = 5$ in (d) and $\rho = 1$ in (e).

Optimal localization radius

- Estimation of ρ_{loc} as a function of σ_{obs} and observation density μ.
- **Approximation error** (or undersampling error[6]) decreases with smaller localization radius.
- **Effective observation error** decreases with a larger ρ, as a larger number of observations gives a better statistical estimates.
- This leads to the *error asymptotics*

$$E(\rho) = \alpha\rho^p + \frac{\beta\sigma_{obs}}{\sqrt{\mu}}\rho^{-\frac{d}{2}}, \quad \rho > 0,$$

- The minimum of $E(\rho) \quad \rightarrow \quad \rho_{min} = c\left(\frac{d}{p}\right)^{\frac{2}{d+2p}}$ with $c(\alpha, \beta, p, \sigma_{obs}, \mu)$

[6] Oke et al. 2007

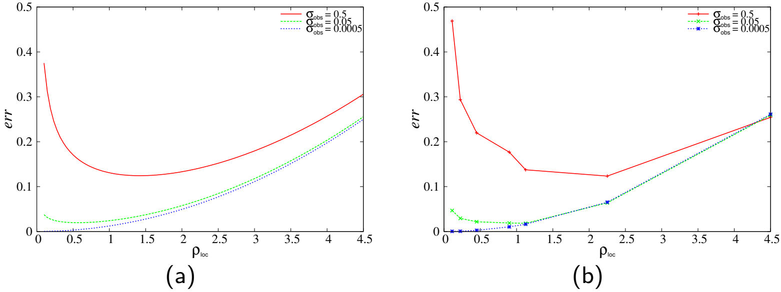
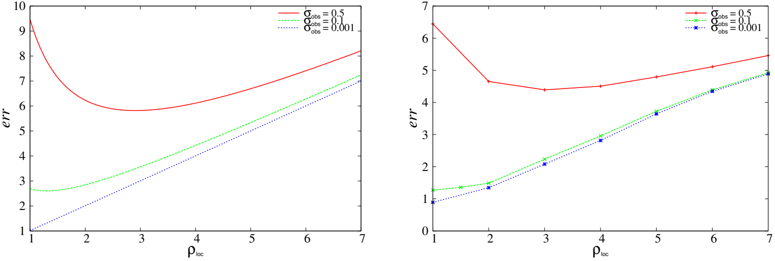


Figure: In (a) we show the theoretical error curve for the case $d = 1$ and $p = 1$. The numerical results (similar curves shown in Greybush et al. 2011) are shown in (b). Here, we display three curves for $\sigma_{obs} \in \{0.0005 \quad 0.05 \quad 0.5\}$.



(a) (b)

Figure: In (a) we show the theoretical error curve for the case $d = 2$ and $p = 1$. Here, we display three curves for $\sigma_{obs} \in \{0.0001 \;\; 0.1 \;\; 0.5\}$.

Outlook / Conclusion

- Optimal localization length ρ_{loc} depending on σ_{obs} and density of observation. These results are analogous for the L95-LETKF.
- For fixed ρ_{loc} in LETKF: $N_{obs} > (N_{ens} - 1)$ gives better results only if ensemble-subspace is appropiated.
- Next steps: Error analysis for **two-step analysis** with different localization radius for each kind of observation assimilated

$$\begin{aligned}
\varphi_1^{(a)} &= \varphi^{(b)} + BH_1^*(R_{1,\rho_1} + H_1 BH_1^*)^{-1}(f_1 - H_1\varphi^{(b)}) \\
\varphi_2^{(a)} &= \varphi_1^{(a)} + B_1 H_2^*(R_{2,\rho_2} + H_2 B_1 H_2^*)^{-1}(f_2 - H_2\varphi_1^{(a)}) \\
\varphi_{total}^{(a)} &= \varphi_2^{(a)}
\end{aligned}$$

- Two-step analysis gives better results if the two observation types have $\sigma_{obs}^1 >> \sigma_{obs}^2$.

Proof of Lemma

Since we have

$$\begin{aligned} \|(f - H\varphi^{(b)}) - HQ_k\gamma\|_{R^{-1}}^2 &= \|(H(\varphi^{(true)} - \varphi^{(b)}) - HQ_k\gamma\|_{R^{-1}}^2 \\ &= \|(\varphi^{(true)} - \varphi^{(b)}) - Q_k\gamma\|_{H^*R^{-1}H}^2 \end{aligned}$$

The element $Q\gamma^{(a)}$ is the *best approximation* in $U^{(L)}$ to the element $\varphi^{(true)} - \varphi^{(b)}$ with respect to the scalar product $\langle \cdot, \cdot \rangle_{H^*R^{-1}H}$.

For the best approximation ψ_* to an element ψ in a Hilbert space X with scalar product $\langle \cdot, \cdot \rangle$ with respect to a finite-dimensional subspace U, for all elements $u \in U$ we have $\langle u, \psi - \psi_* \rangle = 0$.

The analysis $\varphi^{(a)} - \varphi^{(b)} = Q\gamma^{(a)}$ is the orthogonal projection of $\varphi^{(true)} - \varphi^{(b)}$ onto $U_k^{(L)}$.

Finally, we now estimate

$$\begin{aligned} \|\varphi^{(a)} - \varphi^{(true)}\|_{H^* R^{-1} H} &= \|\varphi^{(a)} - \varphi^{(b)} + \varphi^{(b)} - \varphi^{(true)}\|_{H^* R^{-1} H} \\ &= \min_{\gamma} \|Q\gamma - (\varphi^{(true)} - \varphi^{(b)})\|_{H^* R^{-1} H} \\ &= d_{H^* R^{-1} H}\Big(U_k^{(L)}, \varphi^{(true)} - \varphi^{(b)}\Big), \end{aligned}$$

which completes the proof. □

Lemma

Let ψ be a continuously differentiable function defined in some domain V in $\mathbb{R}^n$, $n \in \mathbb{N}$, which satisfies $\psi(x_0) = 0$ for some $x_0 \in V$. Then we have

$$\sup_{x \in B_\rho(x_0)} |\psi(x)| \leq C\rho$$

with the constant $C = \sup_{x \in V} |\nabla \psi(x)|$. In particular, we obtain

$$\sup_{x \in B_\rho(x_0)} |\psi(x)| \ \to 0, \ \ \rho \to 0$$

Proof of Theorem

We carry out the proof for one $x_0 \in V$. We can choose an ensemble member l such that the previous Lemma is satisfied. We set

$$\psi(x) := \frac{|\varphi^{(true)}(x_0) - \varphi^{(b)}(x_0)|}{|\varphi^{(l)}(x_0) - \varphi^{(b)}(x_0|} \left(\varphi^{(l)}(x) - \varphi^{(b)}(x)\right) - \left(\varphi^{(true)}(x) - \varphi^{(b)}(x)\right).$$

yield

$$\sup_{x \in B_\rho(x_0)} |\psi(x)| \leq C\rho \to 0, \quad \rho \to 0.$$

Using the estimate

$$\langle \psi, H^* R^{-1} H\psi \rangle_{B_\rho(x_0)} \leq \|H^*\| \ \|H\| \ \left\|R^{-1}\right\| \langle \psi, \psi \rangle_{B_\rho(x_0)}$$
$$\leq \ \ \|H^*\| \ \|H\| \ \left\|R^{-1}\right\| \ |B_\rho(x_0)| \ \|\psi\|^2_{\infty, B_\rho(x_0)} .$$

Hence, as $\|\psi\|_{\infty, B_\rho(x_0)} = \sup_{x \in B_\rho(x_0)} |\psi(x)| \leq C\rho \to 0, \ \ \rho \to 0.$

This leads to the estimate

$$\|\psi\|^2_{H^* R^{-1} H, \rho} \leq \tau_{H,R} \ \rho \to 0, \ \ \rho \to 0$$

with $\tau_{H,R} := C \|H^*\| \ \|H\| \ \left\|R^{-1}\right\| \ |B_\rho(x_0)|$.

Being ψ the difference between an element in the ensemble space $U^{(L)}$ and $\varphi^{(true)} - \varphi^{(b)}$, we can write

$$\left\| \hat{\varphi}^{(a,\rho)} - \varphi^{(true)} \right\|_{H^* R^{-1} H, \rho} = \left\| \left(\hat{\varphi}^{(a,\rho)} - \varphi^{(b)} \right) - \left(\varphi^{(true)} - \varphi^{(b)} \right) \right\|_{H^* R^{-1} H, \rho}$$

Finally, with the division by $|B_\rho(x_0)|$ now leads to $\hat{E}$ with $\tilde{C} = C \|H^*\| \, \|H\| \, \left\|R^{-1}\right\|$, and the proof is complete. □